NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

THE APPLICABILITY OF NEURAL NETWORKS TO IONOSPHERIC MODELING IN SUPPORT OF RELOCATABLE OVER-THE-HORIZON RADAR

by

James A. Pinkepank

September 1994

Thesis Advisor: Daniel J. Collins

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704 0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
September 1994

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
THE APPLICABILITY OF NEURAL NETWORKS TO IONOSPHERIC MODELING IN SUPPORT OF RELOCATABLE OVER-THE-HORIZON RADAR (U)

5. FUNDING NUMBERS

6. AUTHOR(S)
Pinkepank, James A.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13 ABSTRACT *(Maximum 200 words)*

Ionospheric models have been developed to interpret Relocatable Over-the-Horizon Radar data. This thesis examines the applicability of neural networks to ionospheric modeling in support of Relocatable Over-the-Horizon Radar. Two neural networks were used for this investigation. The first network was trained and tested on experimental ionospheric sounding data. Results showed neural networks are excellent at modeling ionospheric data for a given day. The second network was trained on ionospheric models and tested on experimental data. Results showed neural networks are able to learn many ionospheric models and the modeling network generally agreed with the experimental data.

14. SUBJECT TERMS
Ionosphere Research, Ionospheric Forecasting, Ionospheric Radio Wave Propagation, Neural Networks, Over-the-Horizon Radar

15. NUMBER OF PAGES
52

16. PRICE CODE

17 SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19 SECURITY CLASSIFICATION OF ABSTRACT	20 LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01 280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18

Approved for public release; distribution is unlimited.

THE APPLICABILITY OF NEURAL NETWORKS TO IONOSPHERIC MODELING IN SUPPORT OF RELOCATABLE OVER-THE-HORIZON RADAR

James Alan Pinkepank
Lieutenant, United States Navy
B.S.E.E., University [illegible]a, 1986

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN AERONAUTICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
September 1994

Author: J. A. Pinkepank

Approved by: D. J. Collins, Thesis Advisor

R. E. Ball, Second Reader

D. J. Collins, Chairman
Department of Aeronautics and Astronautics

ABSTRACT

Ionospheric models have been developed to interpret Relocatable Over-the-Horizon Radar data. This thesis examines the applicability of neural networks to ionospheric modeling in support of Relocatable Over-the-Horizon Radar. Two neural networks were used for this investigation. The first network was trained and tested on experimental ionospheric sounding data. Results showed neural networks are excellent at modeling ionospheric data for a given day. The second network was trained on ionospheric models and tested on experimental data. Results showed neural networks are able to learn many ionospheric models and the modeling network generally agreed with the experimental data.

TABLE OF CONTENTS

ACKNOWLEDGMENT

The author would like to thank the Raytheon Company for providing the data package used in this thesis. Special thanks go to Professor Collins for his guidance and patience during the work performed in this investigation.

To my wife, Randi, and my children, Mark and Janice, in appreciation of their continuing love, understanding, patience, and inspiration to complete this work.

I. INTRODUCTION

Relocatable Over-the-Horizon Radar (ROTHR) was developed to support Navy fleet commanders' air defense mission. It was designed to provide air surveillance and warning of attacks by long-range aircraft (primarily bombers) on Navy battle groups and other U.S. and allied tactical forces. (GAO, 1991)

ROTHR is a relocatable, ground-based system with separate transmitter and receiver sites. The transmitters send high frequency signals (5-28 MHz) into the ionosphere that are then refracted downward and reflected off aircraft and other objects. The reflected signals return via the ionosphere to the approximately 8,000 foot receive radar antenna and are processed by computers for target display. ROTHR provides wide-area radar coverage that extends from 500-1,600 nautical miles with a 64-degree azimuth. (GAO, 1991)

The *ionosphere* is the part of the atmosphere that contains enough ions and free electrons to affect radio wave propagation. It starts about 60 km above the earth and extends upward to the atmosphere's outer edge. Reflection off the ionosphere is due to electron interaction with the radio wave electromagnetic fields (Beer, 1976).

A ground-based method of examining the ionosphere is by a sweep frequency sounder known as an *ionosonde*. The ionosonde is a radio transmitter/receiver that transmits a pulse nearly vertically through the atmosphere such that the pulse is reflected off the ionosphere. The frequency of the pulse is altered smoothly and the echo time is recorded as a function of frequency (Ratcliffe, 1972). An *ionogram* plots the echo time against the frequency. An idealized ionogram is shown in Figure 1.1. By knowing the signal travel time and the estimated speed of the pulse, the height of the reflecting layer may be determined.

Three ionospheric *layers* appear quite regularly. The E layer, at about 120 km, is lowest. The F_1 layer, at about 150 to 200 km, is next. Finally, the F_2 layer, at around 250 to 300 km, is the highest layer. (Craig, 1968)

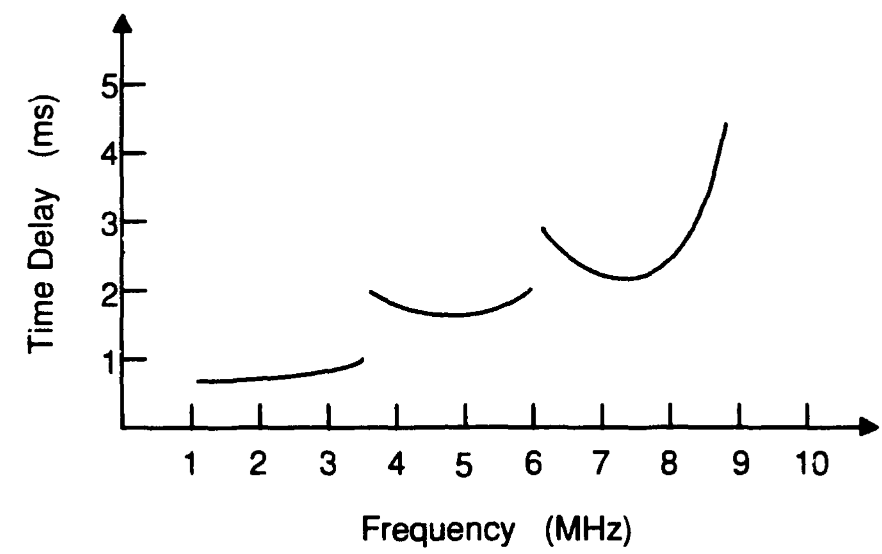

Figure 1.1 Idealized Ionogram. (After Craig, 1968.)

As the pulse's frequency is increased, the reflection altitude increases until reaching a frequency just sufficient for reflection. No reflection occurs for higher frequencies. That is the layer's *critical frequency*. Critical frequencies show up as ionogram discontinuities (see Figure 1.1) and are typically observed at two or more frequencies. (Craig, 1965)

Over 10,000 ionospheric *models* have been developed by the Raytheon Company to interpret ROTHR data. Each ROTHR model is uniquely defined by four numbers: the critical frequencies of the E, F_1, and F_2 layers and the true height of the F_2 layer's peak electron concentration. Proper model selection is a difficult task that requires operator involvement. The present system finds the model that most closely matches the actual sounding. Then the operator has the ability to select an alternate model that may actually be a better match. Operator selection of alternate models requires a well trained operator and this alternate selection process can lead to problems in ROTHR implementation.

This thesis will investigate the application of neural networks to ionospheric modeling. In Chapter II, basic neural network theory is presented and the backpropagation network is introduced. Chapter III describes the procedures used in this investigation and Chapter IV gives the results. Chapter V closes with concluding remarks.

II. NEURAL NETWORKS

A neural network is a nonalgorithmic, nondigital, and intensely parallel distributed information processing system. It consists of a number of relatively simple and highly interconnected processors called *processing elements*. The processing elements are connected by a series of weighted links, over which signals can pass. The network is connected to the outside world through input and output elements. Signals that are put into a network pass through the processing elements and generate a response at the network's output. The neural network has the ability to learn from experience and generalize its knowledge from previous examples. (Caudill, 1992)

A. PROCESSING ELEMENT

The processing element, shown in Figure 2.1, is the fundamental unit in a neural network. Typically, a processing element has many inputs and only one output. The input stimuli are modified by connection weights and then summed. An *activation function* modifies the summed input. This activation function can be a threshold function that only outputs information if the internal activity level reaches a certain value, or it can be a continuous function of the summed input. The activation function's output response is transmitted along the processing element's output connection. This output can be connected to other processing element inputs. (NeuralWare, 1993)

B. FEEDFORWARD NETWORK

Processing elements are highly interconnected and grouped into *layers*. When a fully connected *feedforward network*, such as that shown in Figure 2.2, receives an input vector, each processing element in the input layer receives only an element of the input vector. The input layer processing elements then distribute their input vector elements to the hidden layer processing elements. Due to differing connection weights, each hidden layer processing element sees a different input vector. This causes the hidden layer processing elements to produce differing output responses. Again due to differing connection weights, each output layer processing element sees a different hidden layer

output vector. The output layer processing elements produce the neural network's associated response. (Caudill, 1993)

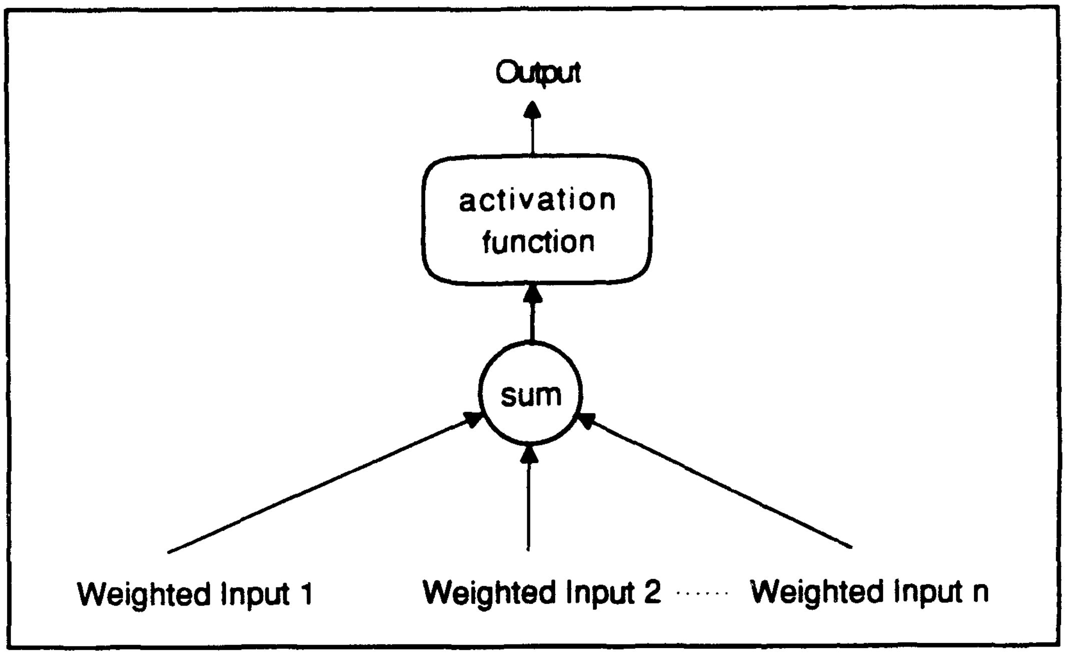

Figure 2.1 Processing Element.

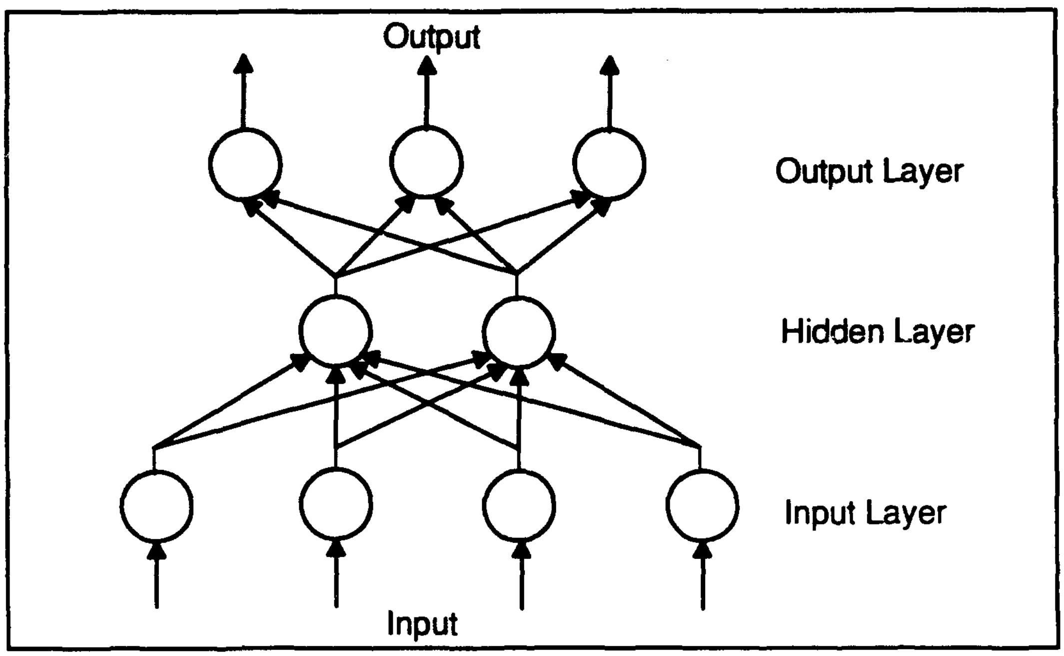

Figure 2.2 Feedforward Neural Network.

C. LEARNING

Neural networks are not programmed like traditional computing systems. They *learn* to solve a problem through training (Caudill, 1993). Learning is achieved through a *learning rule* that systematically changes the connection weights in response to training inputs and optionally the desired outputs of those inputs. The learning rule specifies how connection weights change in response to a training example. A *learning schedule* controls how a learning rule may change over time as the network learns. (NeuralWare, 1993)

Supervised learning occurs when the desired response for each input is presented at the output layer and the network modifies the connection weights to achieve acceptable input/output performance levels. A *hetero-associative* network is a trained network where the desired output is different from the input. (NeuralWare, 1993)

D. MEMORY

The connection weights contain the neural computing memory. The weight values are the current state of network knowledge. An input/output pair is distributed across many memory units in the network and it shares these memory units with other input/output pairs stored in the network. This *distributed* memory characteristic gives the neural network an ability to *generalize*. The network can produce an *intelligent* response when presented with incomplete, noisy, or previously unseen input. (NeuralWare, 1993)

Another distributed memory advantage is neural computing systems are *fault tolerant* and exhibit *graceful degradation*. As processing elements are destroyed or damaged, the network continues to function with only slightly degraded behavior. (NeuralWare, 1993)

E. BACKPROPAGATION

A network with at least one hidden layer must be used to solve complex, non-linearly separable problems. The backpropagation algorithm is a neural network training procedure that provides for hidden layer training. Before the backpropagation training algorithm was developed, neural networks were constrained to one or two layers. A network based on the backpropagation algorithm is an effective multi-layer network that has been extensively used to solve pattern classification problems.

1. Architecture

Typically, a network utilizing backpropagation training is a feedforward network with an input layer, an output layer, and one or more hidden layers. Generally there are no processing element connections within a single layer, and usually each layer is fully connected to the subsequent layer. Research indicates a maximum of three hidden layers are required to solve complex classification problems (NeuralWare, 1993). The hidden layer processing elements act as *feature detectors*. Their connection weights encode the features present in an input. The output layer uses those features to determine the correct response. The ability to generate output based on *features* of the input rather than the raw input data allows the network to create its own complex representation of the problem. (Caudill, 1993)

2. Training and Testing

Backpropagation training is a two-step procedure illustrated in Figure 2.3. First, an input is forward propagated through the network. This causes a response to be generated at the output layer. In the second step, the network's output is compared to the desired output. If the output is not correct, an error signal is generated and passed back through the network with the connection weights being modified as the error backpropagates. (Caudill, 1993)

In testing, an input is presented to the network which generates an output. The output is compared to the desired output and the difference is the test error for that particular example.

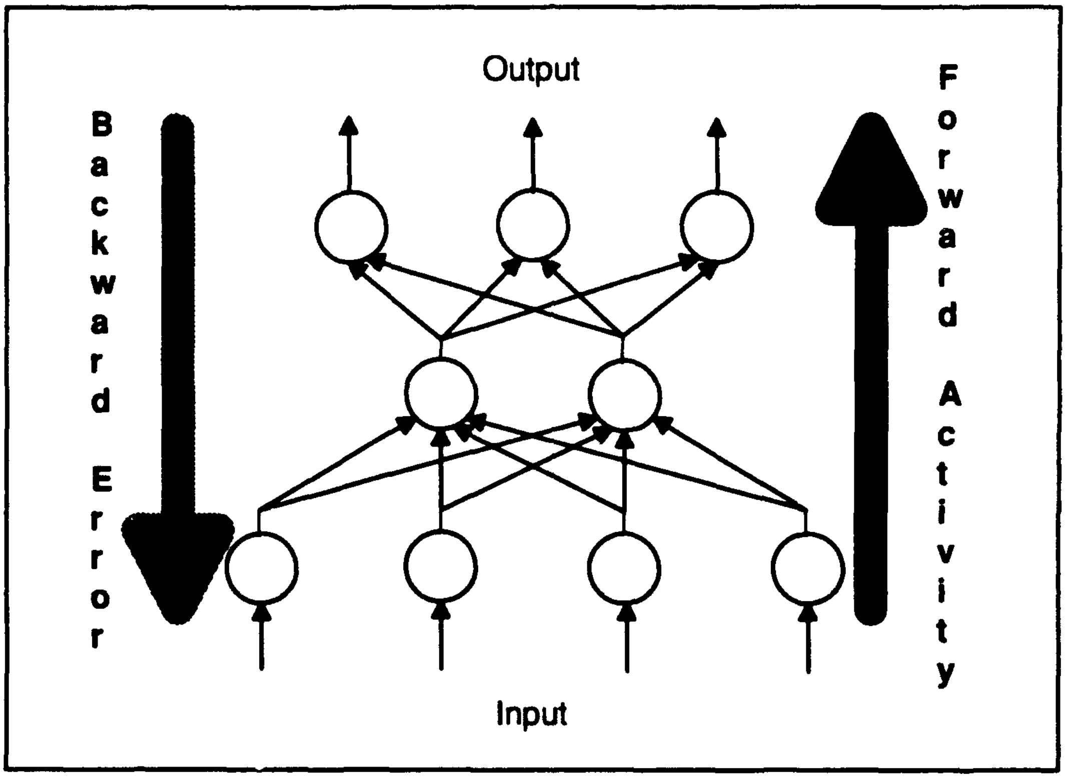

Figure 2.3 Backpropagation Training. (After Caudill, 1993.)

3. Processing Element

A backpropagation processing element's output is determined as follows. First, the weighted input received by processing element j from a total of n processing elements in the network, I_j, is computed:

$$I_j = \sum_{i=1}^{n} w_{ji} x_i \tag{1}$$

The incoming signal from the ith processing element is x_i, and the weight on the connection directed from processing element i to processing element j is w_{ji}. Next, the weighted input passes through the activation function. An activation function commonly used is the sigmoid function (Figure 2.4):

$$f(I) = \frac{1}{1+e^{-I}} \tag{2}$$

The sigmoid function is a *squashing* function with a minimum output value of 0 and a maximum output value of +1. Each processing element's output is usually this activation value. The sigmoid function's derivative is:

$$f'(I) = f(I)(1 - f(I)) \tag{3}$$

The sigmoid function is everywhere differentiable with a positive slope. (Caudill, 1993)

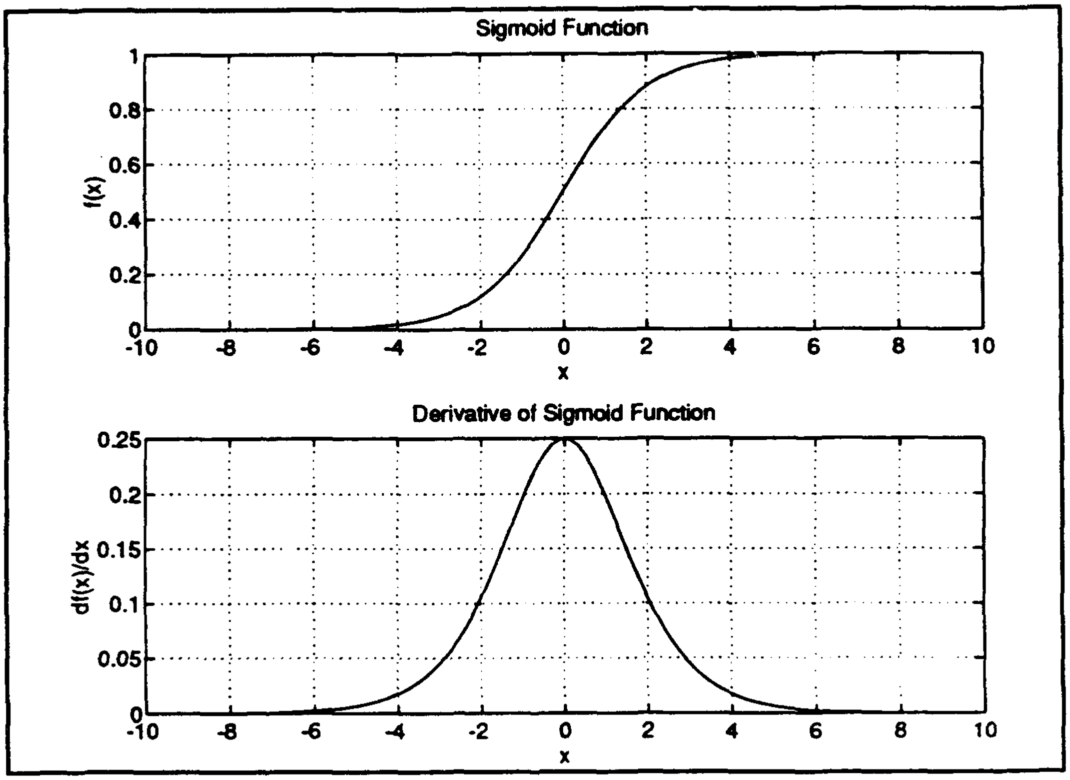

Figure 2.4 Sigmoid Function and Derivative.

4. Learning

The *generalized delta rule* is often used for backpropagation training. The change in a given connection weight is:

$$\Delta w_{ij} = \beta E f(I) \tag{4}$$

where E is the error for this processing element, β is the learning coefficient, a number between zero and one, and $f(I)$ is the processing element input. (Caudill, 1993)

The output layer and hidden layer processing element error terms are computed as follows:

$$E_j^{output} = y_j^{desired} - y_j^{actual} \tag{5}$$

$$E_i^{hidden} = f'(I_i^{hidden}) \sum_{j=1}^{n} (w_{ij} E_j^{output}) \tag{6}$$

Processing element j is in the output layer and processing element i is in a hidden layer. The output layer processing element's output is y. (Caudill, 1993)

The generalized delta rule is a *gradient descent* system that moves the connection weight vector's projection down the steepest descent of the error surface. This is illustrated in Figure 2.5. Multidimensional input and output spaces result in a multidimensional surface instead of the paraboloid shown. (Caudill, 1993)

5. Momentum

A small learning coefficient is desired to avoid divergent behavior but a small learning coefficient leads to very slow learning and a greater possibility of getting stuck in a local minimum. A *momentum* term, α, may be added to the generalized delta rule to resolve this dichotomy:

$$\Delta w_{ij} = \beta E x_i + \alpha \Delta w_{ij}^{previous} \tag{7}$$

β is the learning coefficient, E is the processing element error, and x_i is the processing element input. The momentum term has a value between zero and one. This additional

term allows for faster learning by keeping the weight vector tending to move in the same direction. (Caudill, 1993)

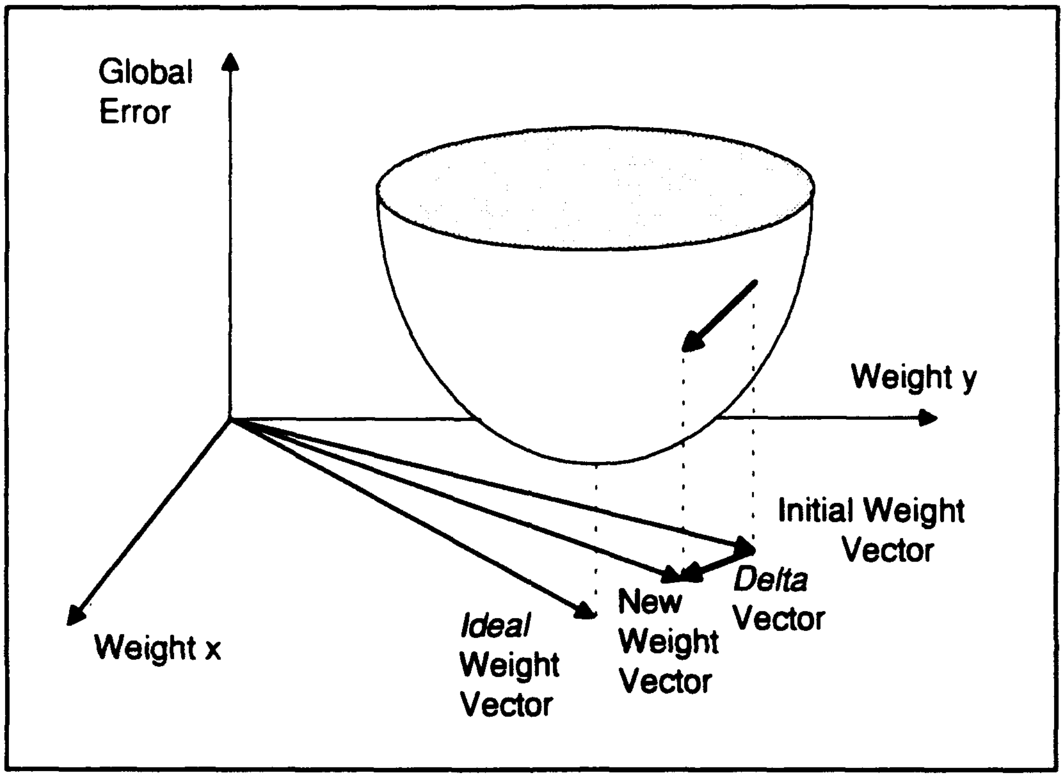

Figure 2.5 Generalized Delta Rule. (After Caudill, 1993.)

III. EXPERIMENTAL PROCEDURE

This chapter discusses the experimental procedures used in this thesis. First to be described is the computing package used in this investigation. Both hardware and software issues are discussed. Then the data package used for this research is described. Data types, structure, and formats are discussed. Finally, a discussion on the two neural networks used in this investigation is presented. Training and test file generation, neural network architecture, and training and testing procedures are all discussed.

A. COMPUTING PACKAGE

Research for this thesis was conducted on a Sun Microsystems, Inc. SPARC2 workstation using the NeuralWare, Inc. NeuralWorks Professional II/PLUS (version 5.0) software package. The MathWorks, Inc. MATLAB (version 4.1) software package was also extensively used.

1. Hardware

The workstation was configured with 64 megabytes of random access memory. This large amount of random access memory allowed a complete training file to be loaded into memory. Loading the entire training file into memory significantly increased I/O speed and saved the hard drive from excessive use (NeuralWare, 1993). The complete ionospheric data package was able to be stored on the workstation's large 2.2 gigabyte hard drive.

2. Software

The Sun OpenWindows workspace provided a multitasking, windowed graphical user interface on top of the SunOS operating system. SunOS is the version of the UNIX operating system used by the workstation. This provided for the ability to simultaneously train multiple networks while performing other data manipulation. (Sun Microsystems, 1991)

NeuralWorks Professional II/PLUS is a multi-model neural network prototyping and development system. It may be used to design, build, train, test, and deploy neural networks to solve complex real-world problems. NeuralWorks has over two dozen well known, built-in network types that can be quickly generated. It also provides for custom network creation. Networks are displayed graphically in full color or monochrome. Network performance may be monitored through an extensive instrumentation package. There are dozens of activation functions and learning rules available. Data for networks can come from the keyboard or an ASCII file. Fully trained feedforward networks may be converted into C code providing a built-in facility for deploying developed networks. NeuralWorks Professional II/PLUS is a very powerful neural network development system. (NeuralWare, 1993)

MATLAB is another software package used in this research. It was used for numeric computation, data manipulation, and graphing. MATLAB is a technical computing environment written in C code for high-performance numeric computation and visualization (MathWorks, 1992).

B. DATA PACKAGE

The Raytheon Company provided the data package used in this investigation. It consisted of a *Quasi-Vertical-Incidence* (QVI) sounding data tape, the ROTHR model QVI library data tape, and a computer printout that shows the QVI model that the current pattern recognition algorithm chose to best-fit each QVI sounding as modified by an *expert* observer.

1. QVI Sounding Data

A Sun workstation compatible data tape contained *grayscale* and *peak* QVI ionospheric soundings for a 24 hour period on 3 May 1990. The soundings were recorded every 10 minutes.

The grayscale data is the raw sounding information (a two-dimensional array giving received power as a function of sounding frequency and time delay). The peak data is an abstracted version of the grayscale data, in which the two-dimensional array has been searched to find points which have great enough signal-to-noise ratio to probably be real returns and which are local peaks in range and frequency. The intent of converting the grayscale data to the peak data is to reduce the real-time computational load on the ROTHR data processing equipment. (Thome, 1991)

Figures 3.1-3.3 show the first three QVI peak soundings recorded.

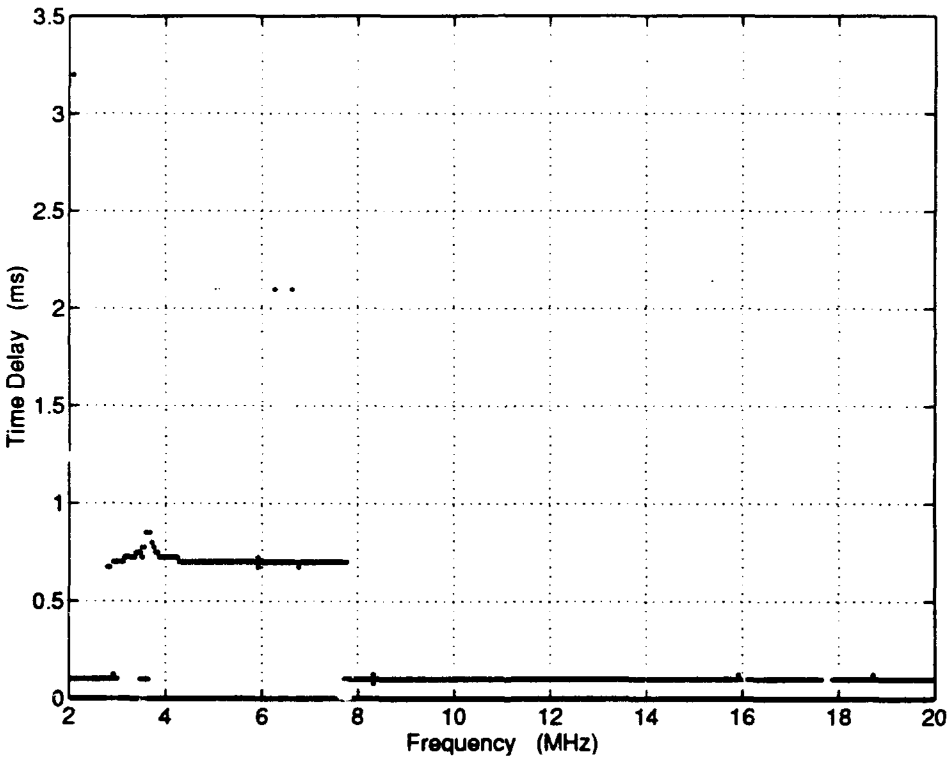

Figure 3.1 QVI Peak Sounding, 3 May 1990, 0008Z.

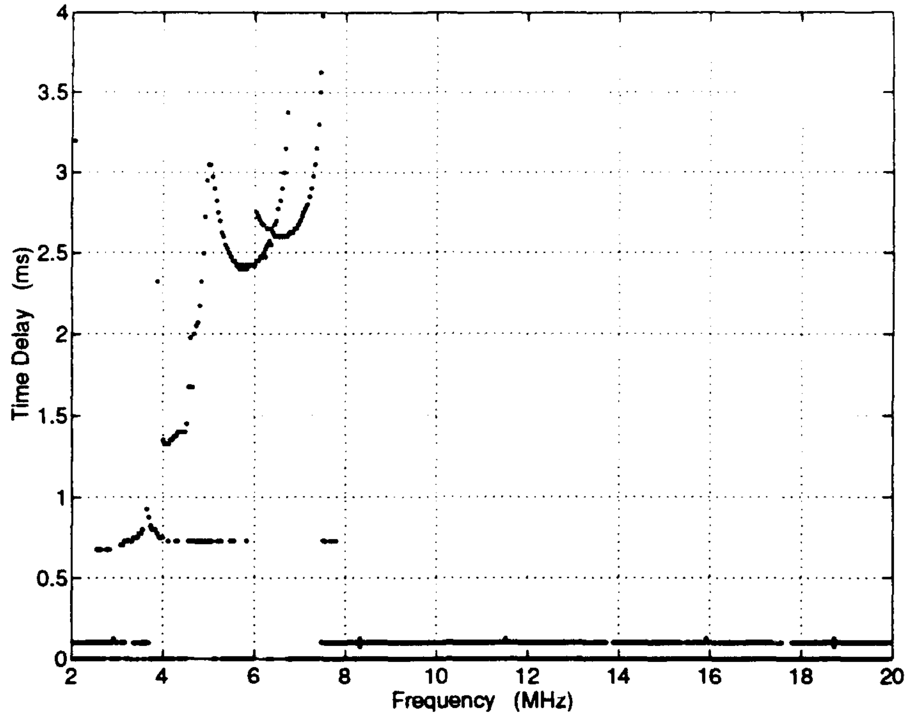

Figure 3.2 QVI Peak Sounding, 3 May 1990, 0018Z.

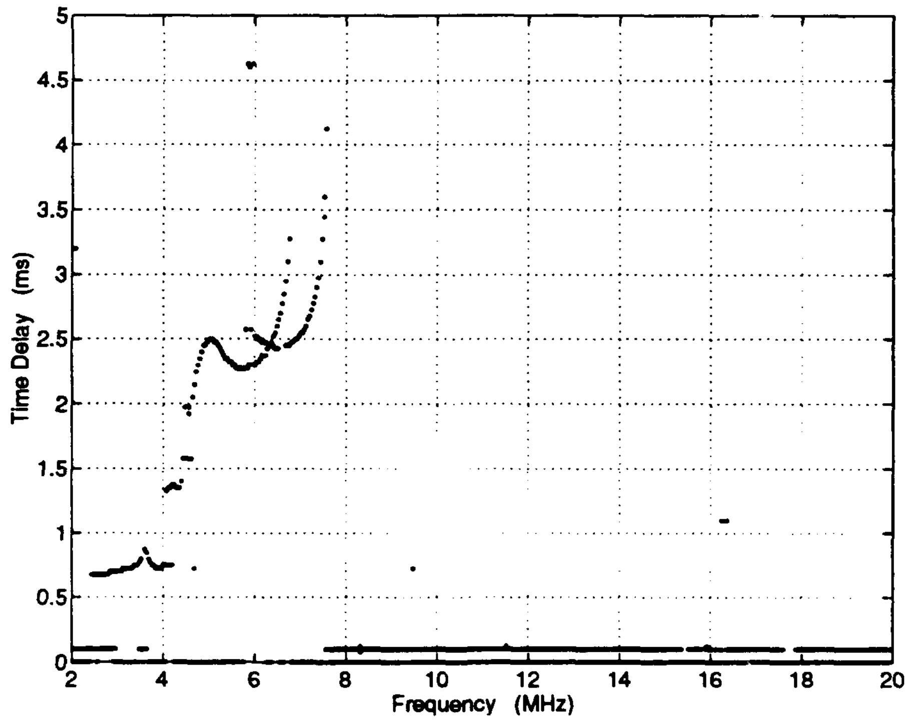

Figure 3.3 QVI Peak Sounding, 3 May 1990, 0028Z.

2. ROTHR QVI Library Data

A Sun workstation compatible data tape contained the ROTHR model QVI library in four files. There are over 10,000 models in this library.

> Each model is uniquely defined by four numbers: the critical frequencies of the E, F_1, and F_2 layers and the true height of the peak of the F_2 layer. For each model contained in the library, there is stored on tape a set of points which define a model QVI sounding (in the same coordinate system and with the same granularity as for the observed QVI soundings). (Thome, 1991)

Figure 3.4 shows a sample QVI library model sounding and Figure 3.5 shows the QVI library model data format.

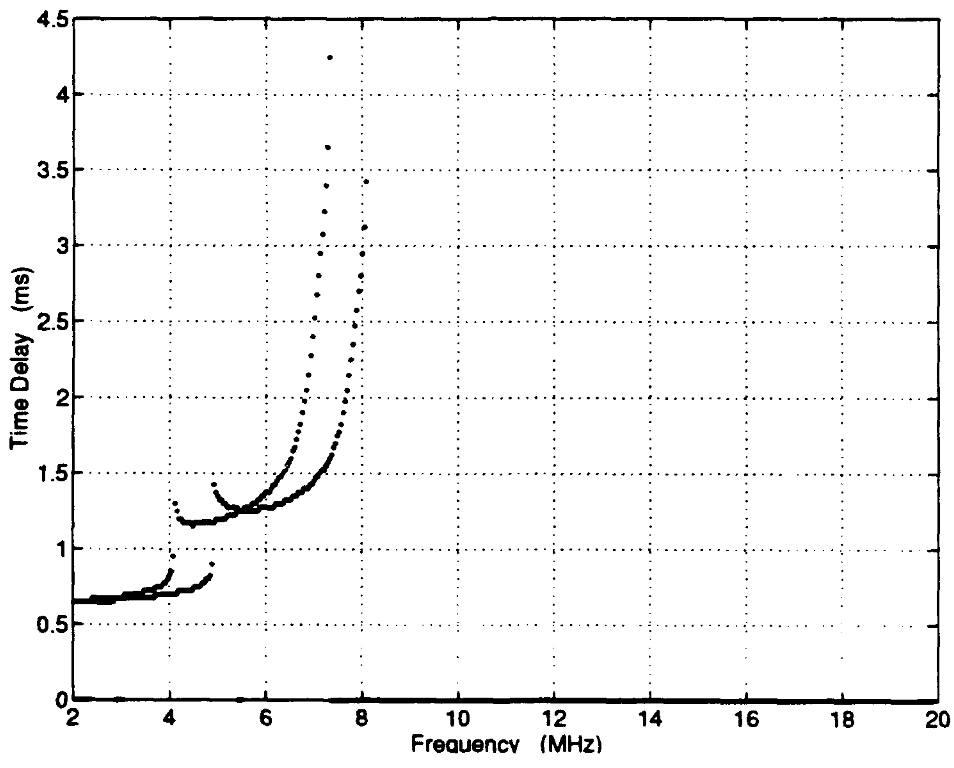

Figure 3.4 Sample QVI Library Model Sounding.

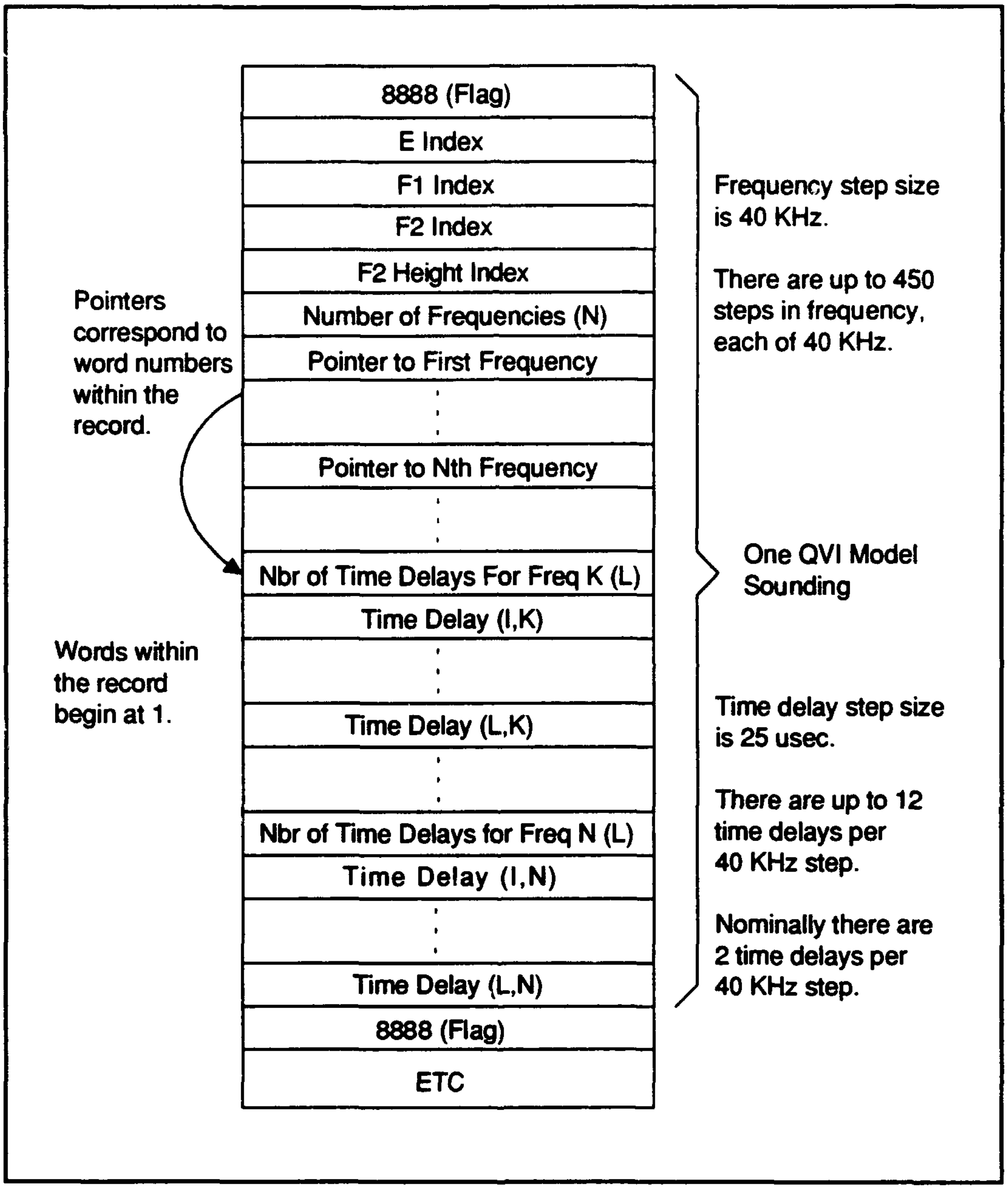

Figure 3.5 QVI Library Model Data Format. (After Thome, 1991.)

3. Expert Data

Expert data came from a computer printout that listed the four model-defining numbers of each QVI model the current pattern matching algorithm (the *expert*) chose to best-fit each observed QVI sounding. Figure 3.6 shows the E layer expert data. Figure 3.7 shows the F_1 layer expert data. Figure 3.8 shows the F_2 layer expert data. Finally, Figure 3.9 shows the F_2 layer peak height expert data. The expert data shows the diurnal variation of the various layers in a discretized manner.

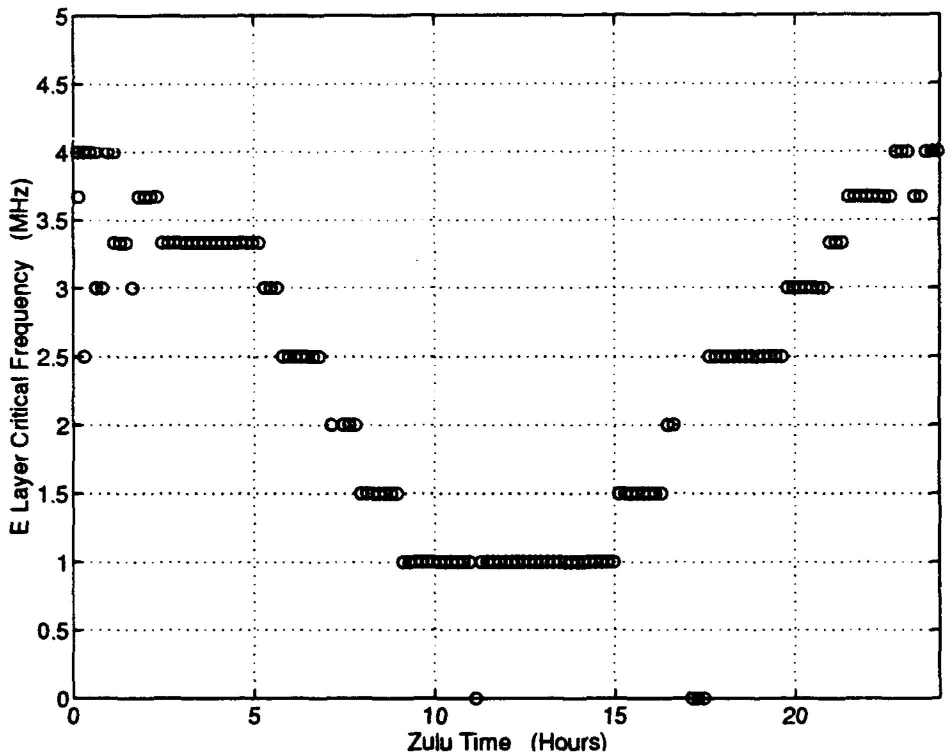

Figure 3.6 E Layer Expert Data, 3 May 1990.

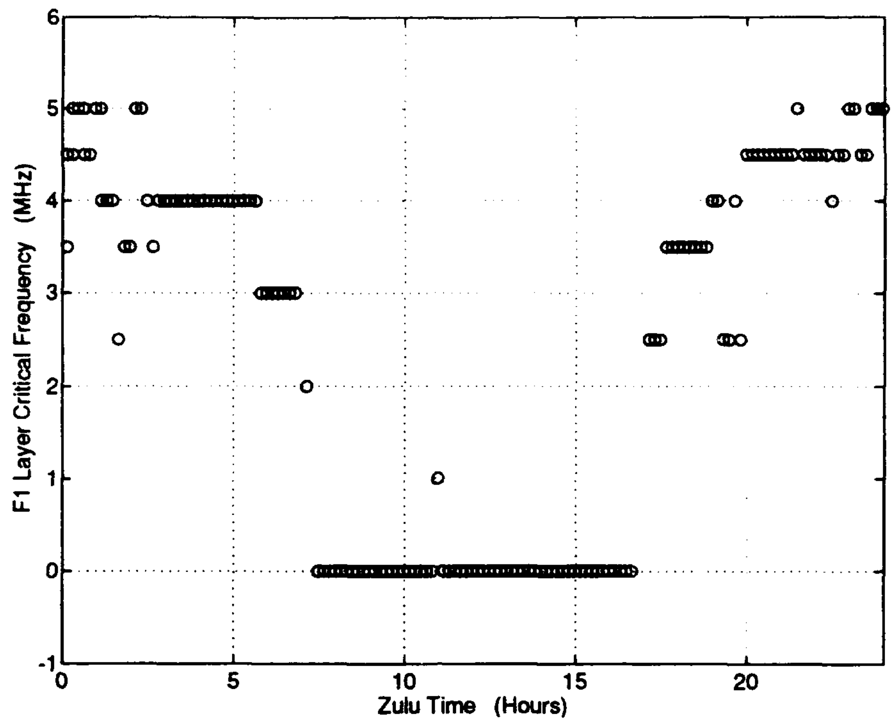

Figure 3.7 F_1 Layer Expert Data, 3 May 1990.

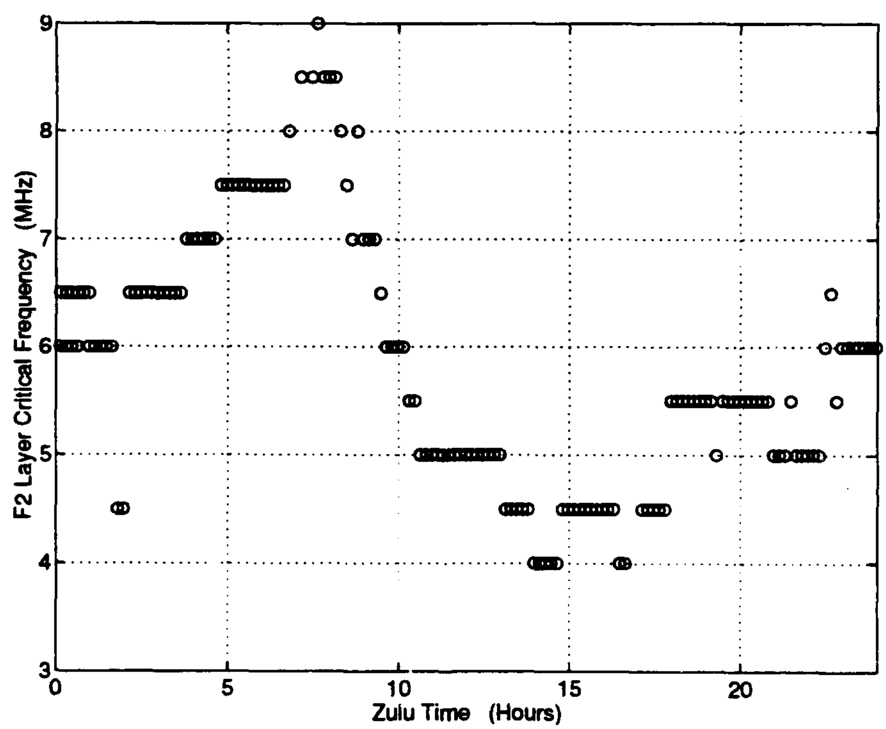

Figure 3.8 F_2 Layer Expert Data, 3 May 1990.

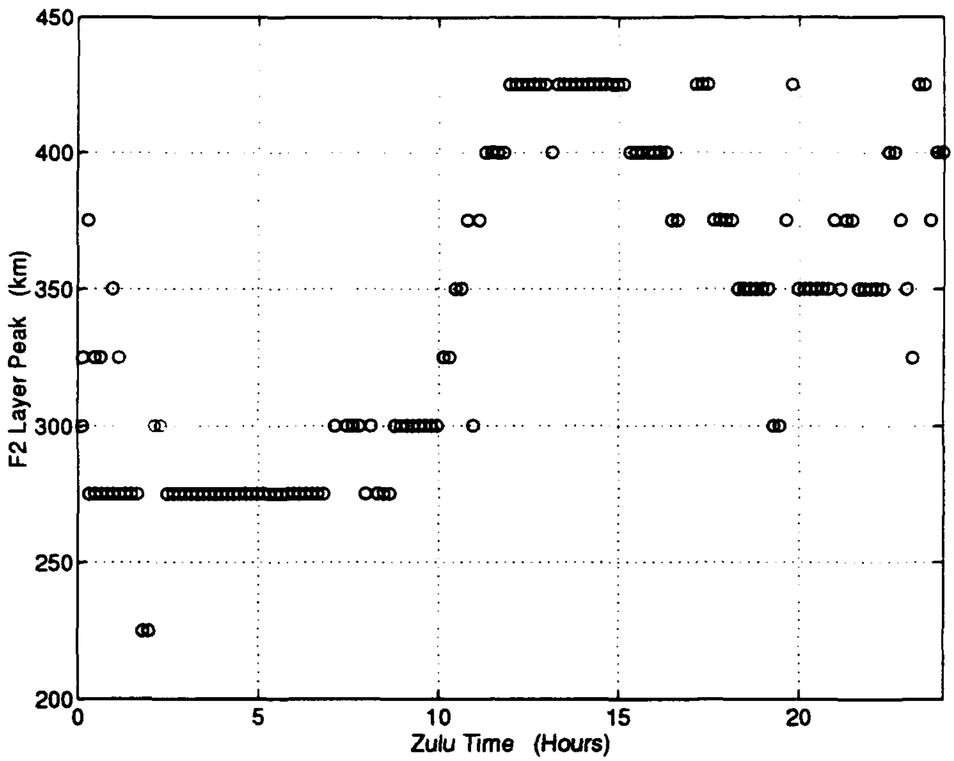

Figure 3.9 F_2 Layer Peak Height Expert Data, 3 May 1990.

C. EXPERIMENTAL SOUNDING NEURAL NETWORK

The experimental sounding neural network was trained on half the experimentally recorded QVI sounding data. Then it was tested on the same training data to test the networks ability to *recall* previously presented examples. Finally, the network was tested on the other half of experimentally recorded QVI sounding data to check independent test data performance. The specific steps required to perform this portion of the research is given in the following paragraphs.

Before constructing the neural network, the training and test files were created. There were 146 QVI peak soundings with corresponding expert data. Every other sounding/expert-data pair was placed in the training file. The other sounding/expert-data pairs were placed in the test file. That resulted in a 74 example training file and a 72 example test file.

Figure 3.6 shows there are four soundings where the expert assigned an E layer critical frequency of 0 MHz. Since the E layer does not normally disappear at night, these values are possibly in error. Therefore, the four soundings in question were removed from the training and test files for the majority of this investigation. However, the effect of including the four questionable soundings in the training and test sets was investigated and the results are reported in Chapter IV.

The revised training file had 72 records and the revised test file had 70 records. Each record consisted of an input/output pair containing 902 items of data. The input was an observed QVI sounding and the output was the sounding's corresponding expert data. The input specifically consisted of 898 time delays (two time delays for each of the first 449 frequencies recorded).

After creating the training and test files, the neural network was constructed with the *Backpropagation* command in the NeuralWorks *InstaNet* menu. The QVI Peak Sounding Neural Network was a fully connected, feedforward, backpropagation network with two hidden layers. There were 898 processing elements in the input layer (one for each time delay), 100 processing elements in the first hidden layer, 20 processing elements in the second hidden layer, and four processing elements in the output layer. The output layer processing elements returned the E, F_1, and F_2 layer critical frequencies as well as the F_2 layer peak height. The processing element activation function was the sigmoid function and the learning rule was the generalized delta rule with momentum. The network was now ready to be trained and then tested.

A problem that can occur with backpropagation networks is the problem of *over training*. Over training a neural network results in some loss of the network's ability to generalize. When over trained, the network performs well on the training data but poorly on independent test data. The problem of over training is handled through the NeuralWorks *SaveBest* training option. SaveBest runs train/test cycles and automatically saves the best performing network based on the performance criteria selected. In this investigation, the performance criteria selected was the Root Mean Square (RMS) error for *all* processing elements in the output layer. (NeuralWare, 1993)

Through the SaveBest training option, the network was trained 100,000 times on the training file examples and tested every 1,000 training iterations on the test file examples. The training rate was approximately 11,000 examples/hour. Examples were presented until all examples were used once. Then another random pass was made through the data set, etc. This continued until 100,000 examples were presented to the network. The results are presented in Chapter IV.

D. MODEL NEURAL NETWORK

The model neural network was trained on ROTHR model QVI library data. Then it was tested on *all* experimental data (142 soundings) to see how the model network's output compared to the *expert's* output. The specific steps required to perform this portion of the research is given in the following paragraphs.

The first item to be created for this neural network was the training file. ROTHR QVI library data was contained on four files that were grouped by the range of F_2 values they contained. Figure 3.8 shows an F_2 layer critical frequency range of 4-9 MHz for the 24 hour period. This range of F_2 values was chosen as the training set range. There were 6,878 library models available with an F_2 layer critical frequency in the 4-9 MHz range. Every other model in that range was placed in the training file resulting in 3,439 examples. A training example's input consisted of 400 model time delays (two time delays for each frequency below 10 MHz) and its output was the model's four defining parameters. The frequency range was decreased for this network to aid in reducing the training time required.

Next, the test file was created. The test file consisted of all 142 QVI peak soundings and corresponding expert data that were previously used for the QVI Peak Sounding training and test files. A test record's input consisted of 400 measured time delays (two time delays for each frequency below 10 MHz) and its output was the sounding's corresponding expert data.

After creating the training and test files, the neural network was constructed with the Backpropagation command in the NeuralWorks InstaNet menu. The ROTHR QVI

Library Neural Network is a fully connected, feedforward, backpropagation network with two hidden layers. There are 400 processing elements in the input layer, (one for each time delay), 100 processing elements in the first hidden layer, 20 processing elements in the second hidden layer, and four processing elements in the output layer. The output layer processing elements returned the E, F_1, and F_2 layer critical frequencies as well as the F_2 layer peak height. The processing element activation function was the sigmoid function and the learning rule was the generalized delta rule with momentum. Now the network was ready for training and testing.

Through the SaveBest training option, the network was trained 2,310,000 times on the 3,439 training file examples and tested every 1,000 training iterations on the 142 test file examples. The training rate was approximately 20,000 examples/hour. Because there were more than 2,500 training examples, NeuralWorks simply randomly selected training examples until 2,310,000 examples were presented to the network. The results are presented in Chapter IV.

IV. RESULTS

This chapter presents the results found in this investigation. The experimental sounding neural network's results are discussed first followed by the model neural network's results.

A. EXPERIMENTAL SOUNDING NEURAL NETWORK

Figure 4.1 plots the experimental sounding network's output layer *test set* RMS error as a function of training received. The optimal amount of training, defined as the smallest output layer RMS error for the experimental sounding test set, is 5,000 training iterations. Until 5,000 iterations, the network's test performance steadily improved. Between 5,000 and 40,000 iterations, the network's test performance generally declined. After 40,000 iterations, the network had essentially memorized the training set and little additional learning was taking place.

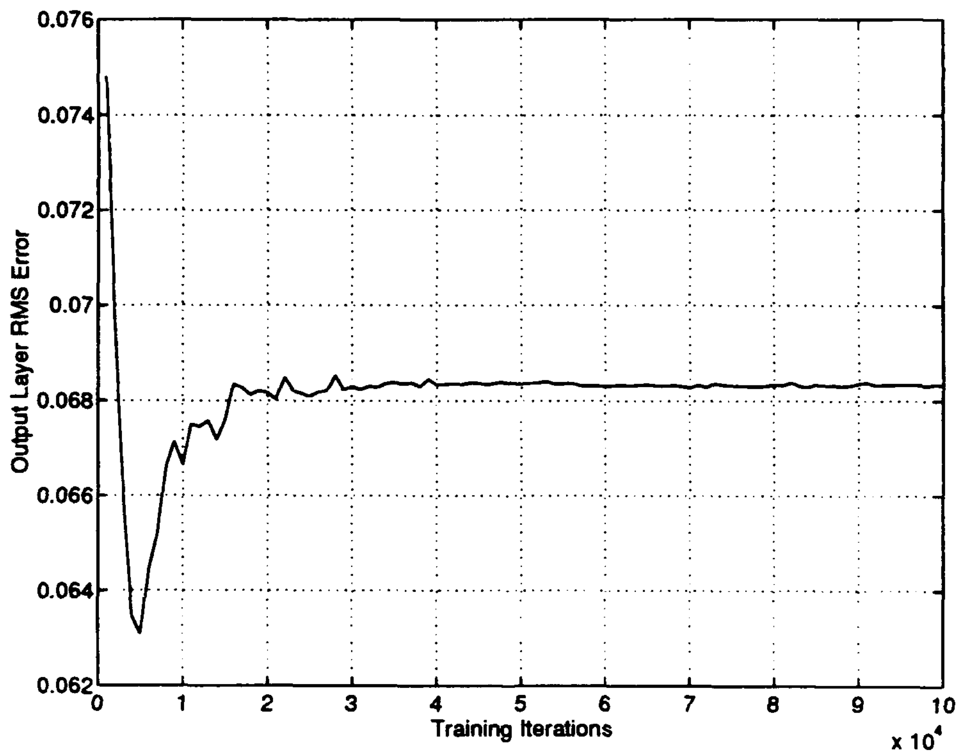

Figure 4.1 Test Set Error for the Experimental Sounding Neural Network.

Figures 4.2 and 4.3 contrast the experimental sounding network's E layer *train set* results for the optimally trained and over trained networks. The optimally trained network (Figure 4.2) has learned the layer's diurnal variation with a train set critical frequency RMS error of 0.1549 MHz. The over trained network (Figure 4.3), with an RMS error of 0.0011 MHz, has essentially memorized the train set.

Figures 4.4 and 4.5 contrast the experimental sounding network's E layer *test set* results for the optimally trained and over trained networks. The optimally trained network (Figure 4.4) has learned the layer's diurnal variation with a test set critical frequency RMS error of 0.3293 MHz. The over trained network (Figure 4.5) exhibits a larger test set RMS error of 0.3453 MHz. This larger error is due to over training that has degraded the network's ability to generalize.

The optimally trained network similarly exhibited superior performance on F_1, F_2, and F_2 layer peak test data. Therefore, only the optimally trained network's test results for the F_1, F_2, and F_2 layer peak will be discussed.

Figure 4.6 shows the experimental sounding network's F_1 layer *test set* results. The optimally trained network has learned the layer's diurnal variation with a test set critical frequency RMS error of 0.6705 MHz. Because the *train set* included a probable anomalous reading of 1 MHz at approximately 1100Z (shown in Figure 3.7), the network's *test set* results show a corresponding cluster of responses near 1 MHz around 1100Z.

Figure 4.7 shows the experimental sounding network's F_2 layer *test set* results. The optimally trained network has learned the layer's diurnal variation with a test set critical frequency RMS error of 0.4109 MHz. Because the network was trained on a probable anomalous reading of 4.5 MHz at approximately 0200Z, the network's responses are lower than they should be around 0200Z.

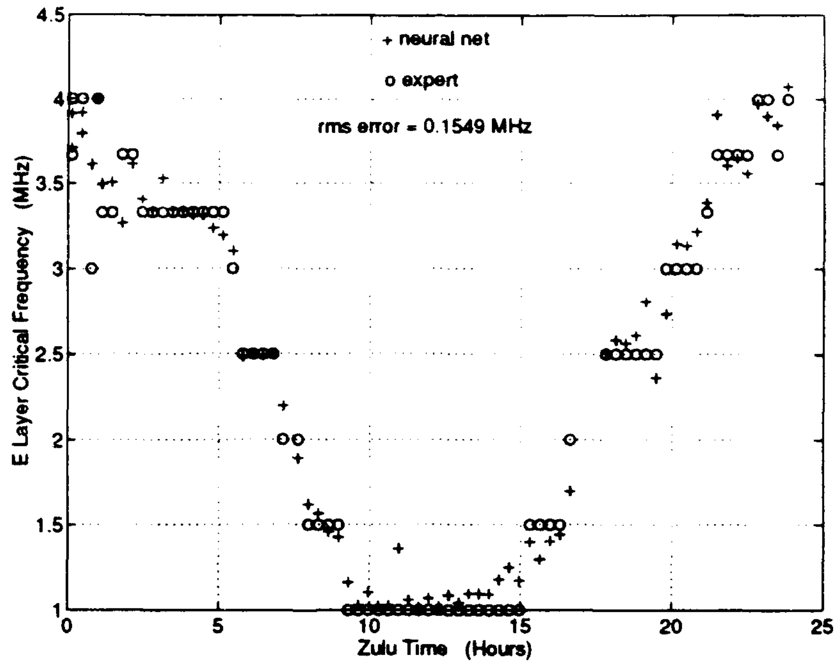

Figure 4.2 Train Set (E Layer) - Experimental Sounding Network. 5,000 training passes.

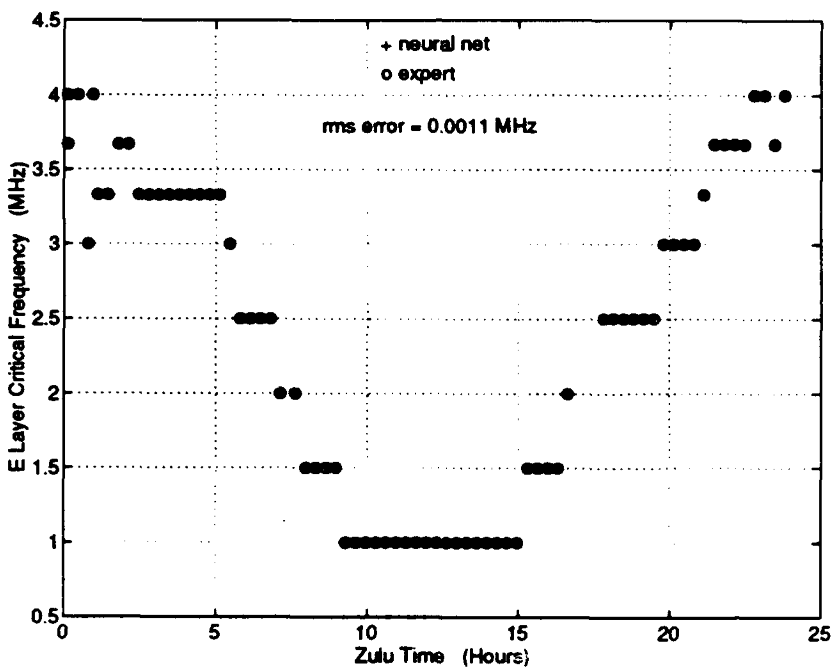

Figure 4.3 Train Set (E Layer) - Experimental Sounding Network. 100,000 training passes.

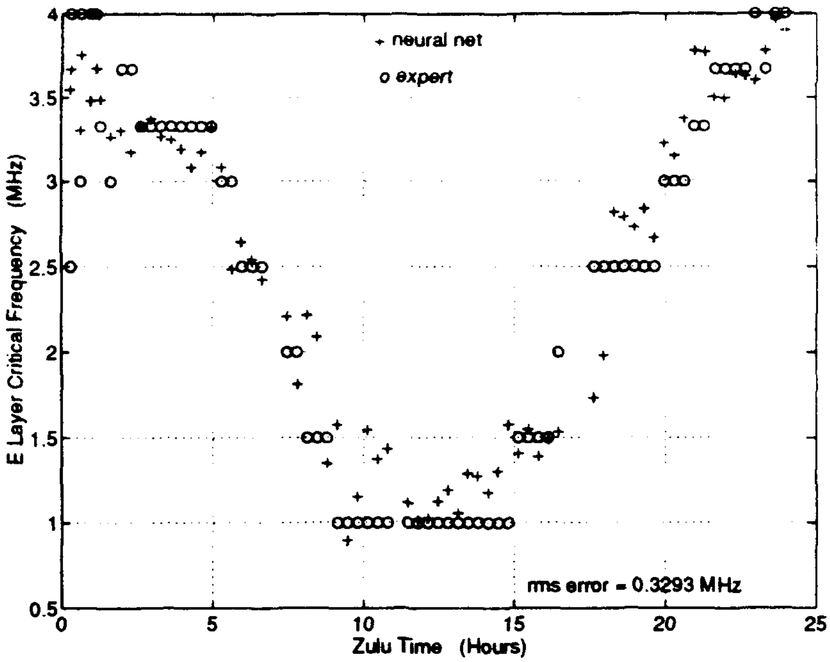

Figure 4.4 Test Set (E Layer) - Experimental Sounding Network. 5,000 training passes.

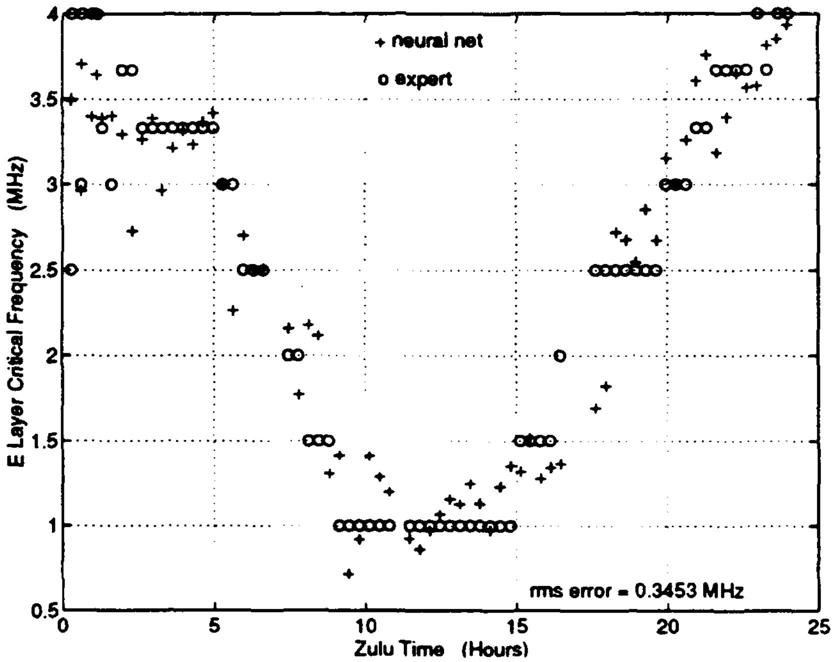

Figure 4.5 Test Set (E Layer) - Experimental Sounding Network. 100,000 training passes.

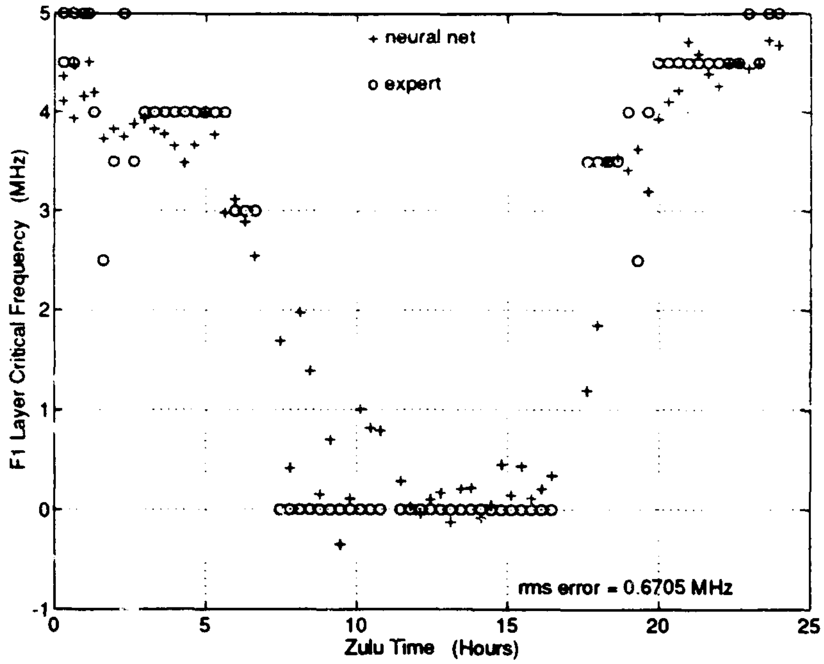

Figure 4.6 Test Set (F_1 Layer) - Experimental Sounding Network. 5,000 training passes.

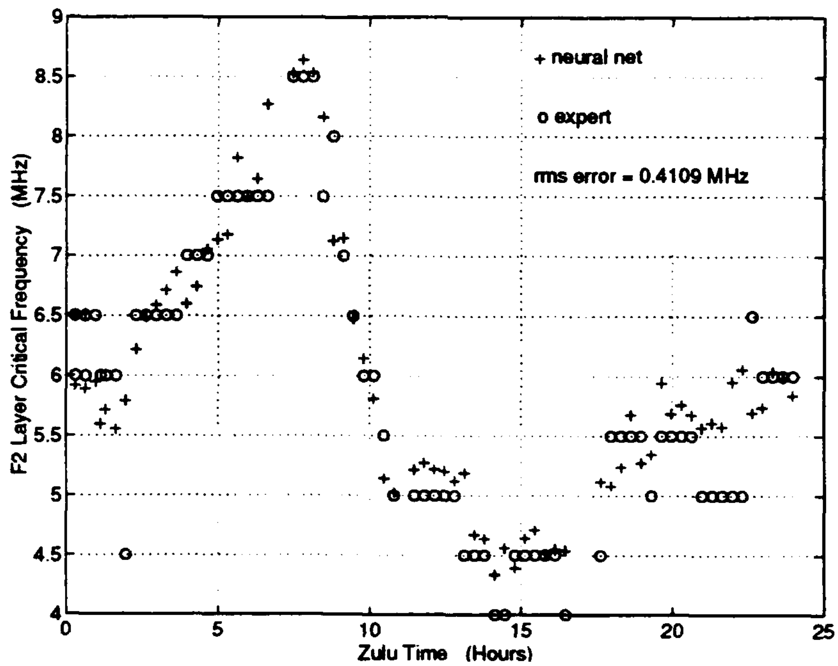

Figure 4.7 Test Set (F_2 Layer) - Experimental Sounding Network. 5,000 training passes.

Figure 4.8 shows the experimental sounding network's F_2 layer peak *test set* results. The optimally trained network has learned the layer's diurnal variation with a test set F_2 layer peak height RMS error of 22.5358 km. The network exhibits declining performance after 1700Z due to widely scattered expert data. Also, the network was trained on three probable anomalous readings that were in the train set data. The anomalies were 225 km at approximately 0200Z, 300 km at approximately 1930Z, and 325 km at approximately 2300Z. The network ignored the outlier at 0200Z, but it did return a cluster of approximately 300 km responses around 1930Z and a cluster of approximately 325 km responses around 2300Z

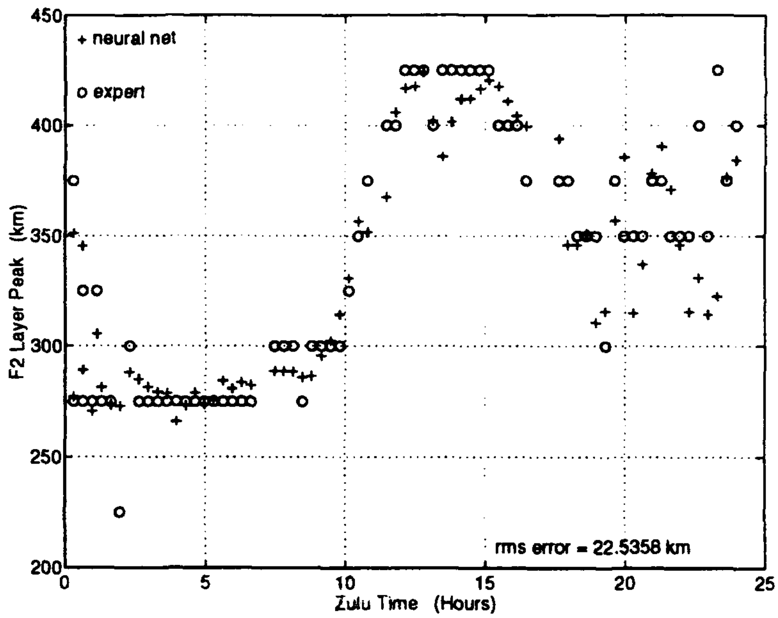

Figure 4.8 Test Set (F_2 Layer Peak) - Experimental Sounding Network. 5,000 training passes.

The effect of what is thought to be anomalous data can be seen in the following comparison. Figure 4.9 shows an experimental sounding network that was trained and tested on data that included suspected anomalous readings at approximately 1100Z and 1700Z. Figure 4.10 shows an experimental sounding network that was trained and tested on data that excluded the anomalous soundings. By comparing the two figures it may be

seen that the anomalous data caused the E layer critical frequency to drop below 1 MHz between 1100Z and 1400Z and it also caused a group of low values around 1700Z. Including the suspected anomalous readings resulted in an 18% larger RMS error in the E layer critical frequency.

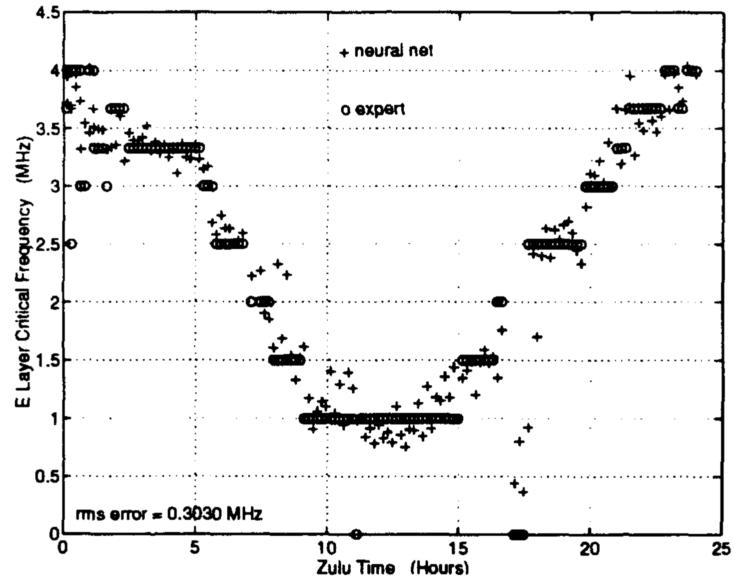

Figure 4.9 Experimental Sounding Network Trained on Anomalous Readings.

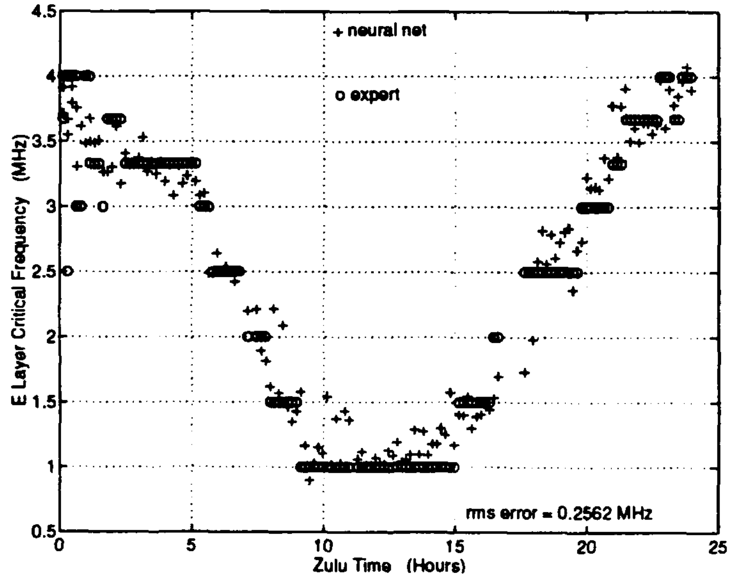

Figure 4.10 Experimental Sounding Network Trained on Reduced Data Set.

B. MODEL NEURAL NETWORK

Figure 4.11 plots the model network's output layer *test set* RMS error as a function of training received. The optimal amount of training, defined as the smallest output layer RMS error for *all* experimental data, is 10,000 training iterations. Until 10,000 training iterations, the network's test performance generally improved. Between 10,000 and 70,000 training iterations, the network's test performance generally declined. Between 70,000 and 1,500,000 training iterations the network's test performance generally improved but at a much slower rate than before. After 1,500,000 training iterations, the network had essentially memorized the training set.

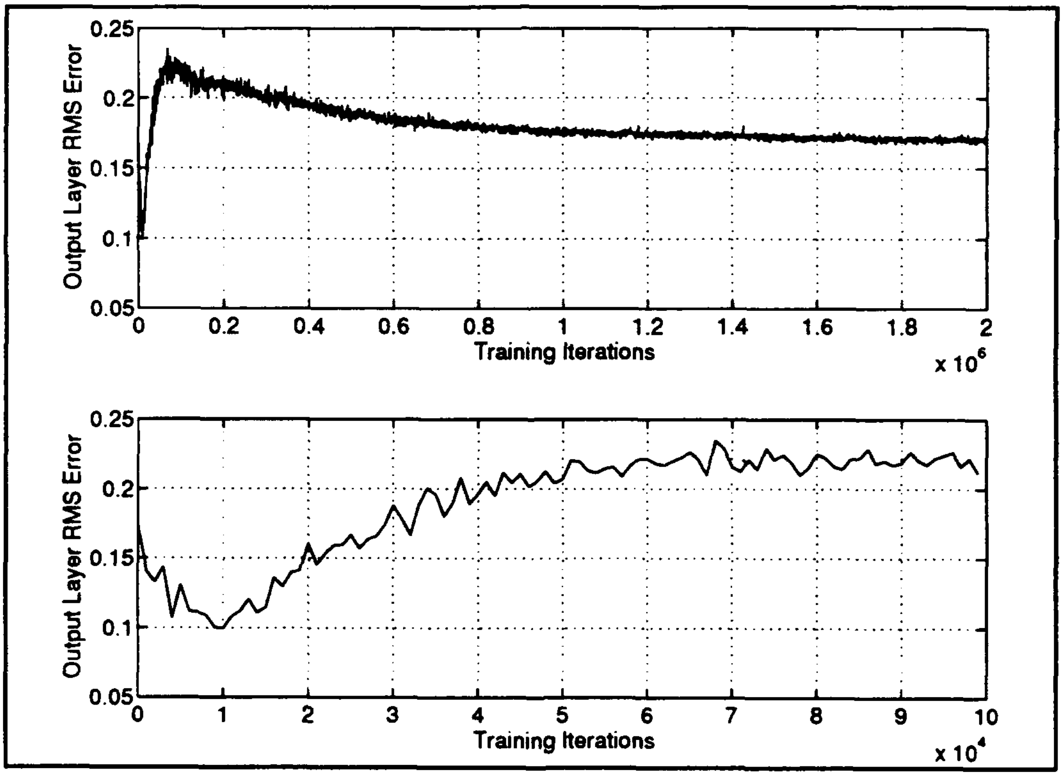

Figure 4.11 Test Set Error for the Model Neural Network.

The lower graph is an expanded portion of the upper graph.

Figures 4.12 and 4.13 contrast the model network's E layer *test set* (all experimental data) results for the optimally trained and over trained networks. The optimally trained network (Figure 4.12) correctly modeled the layer's diurnal variation and had a test set critical frequency RMS error of 0.6665 MHz. This may imply that for this layer the models are very good and the experimental data is accurate. The optimally trained network performed much better than the over trained network on test data. The over trained network (Figure 4.13) exhibits a much larger test set critical frequency RMS error of 1.4188 MHz. This large increase in error is due to over training.

The optimally trained network similarly exhibited superior performance on F_1, F_2, and F_2 layer peak test data. Therefore, only the optimally trained network's test results for the F_1, F_2, and F_2 layer peak will be discussed.

Figure 4.14 shows the model network's F_1 layer *test set* results. The optimally trained network has modeled the layer's general diurnal variation shape but the results do not agree with the experimental data between approximately 0800Z and 1600Z. The network returned values of approximately 1 MHz while the experimental data showed values of 0 MHz. One possible explanation might be that there were not enough training examples that matched the experimentally recorded conditions for that time frame. Only 20 out of 3,439 models in the *train set* had an F_1 layer critical frequency of 0 MHz, an E layer critical frequency of 1 MHz, *and* an F_1 layer peak height above 350 km. Those three conditions were experimentally recorded for the majority of the time between 0800Z and 1600Z. Therefore, the network was not trained much on a pattern that occurred often in the test set.

Figure 4.15 shows the model network's F_2 layer *test set* results. The optimally trained network correctly modeled the layer's diurnal variation and had a test set critical frequency RMS error of 0.5294 MHz. This may imply that for this layer the models are very good and the experimental data is accurate.

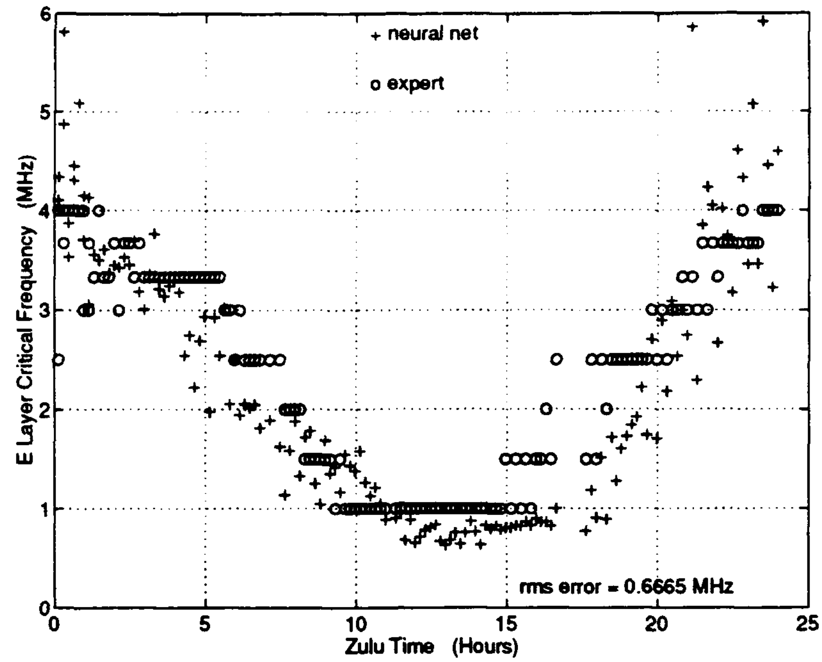

Figure 4.12 Test Set (E Layer) - Model Network.

10,000 training passes.

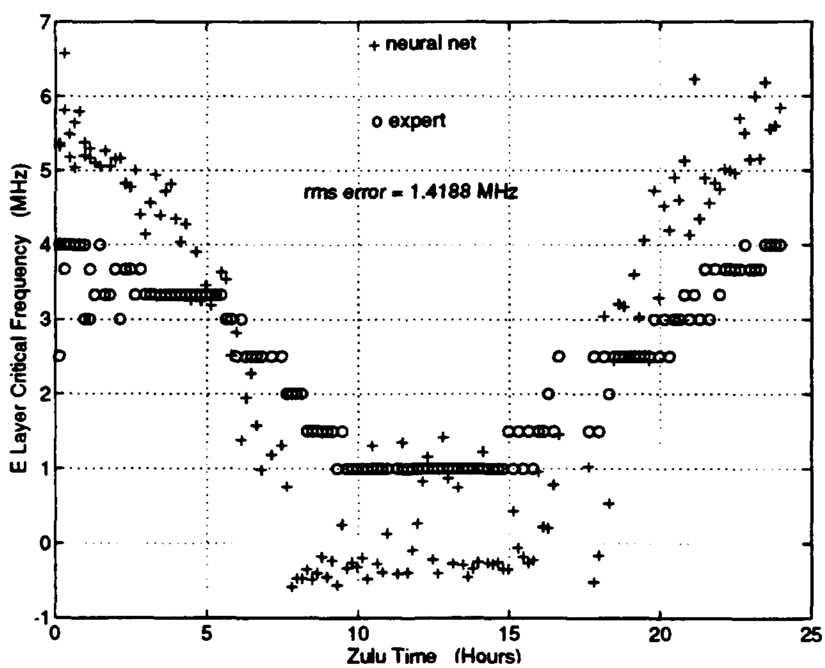

Figure 4.13 Test Set (E Layer) - Model Network.

2,310,000 training passes.

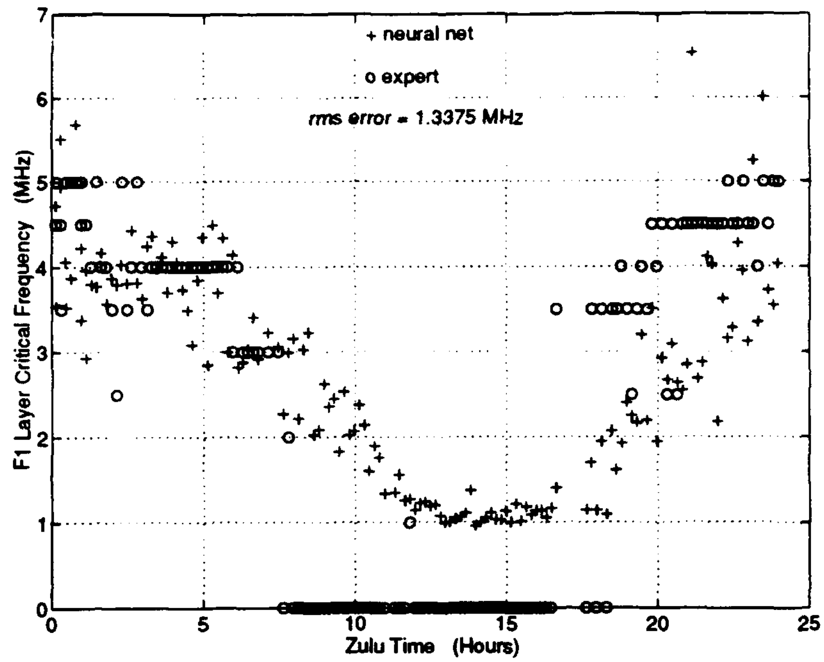

Figure 4.14 Test Set (F_1 Layer) - Model Network. 10,000 training passes.

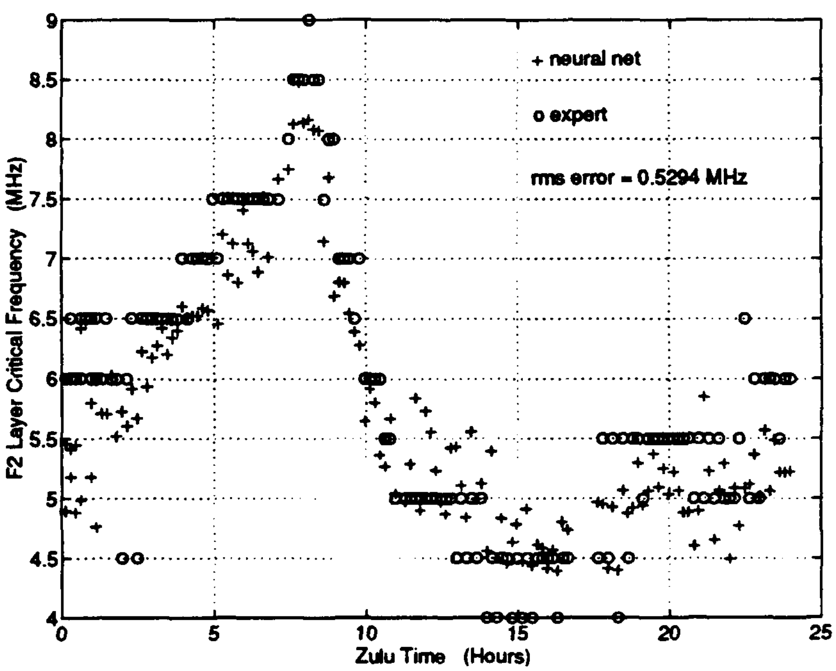

Figure 4.15 Test Set (F_2 Layer) - Model Network. 10,000 training passes.

Figure 4.16 shows the model network's F_2 layer peak *test set* results. The optimally trained network's responses show a general pattern reflecting the daily variation of the F_2 layer peak height but they are widely scattered. Between 1700Z and 2400Z there are large changes in the expert data over short time frames. This may imply there is a large uncertainty in the data during that time period. Therefore, the network's error may be caused by the combination of a large uncertainty in the expert data and the possibility that the modeling of the F_2 layer peak may not be quite adequate. Modeling the F_2 layer peak is complicated by the fact that its behavior is not yet fully understood in detail and must represent the net effect of many individual physical processes (Ivanov-Kholodny, 1986).

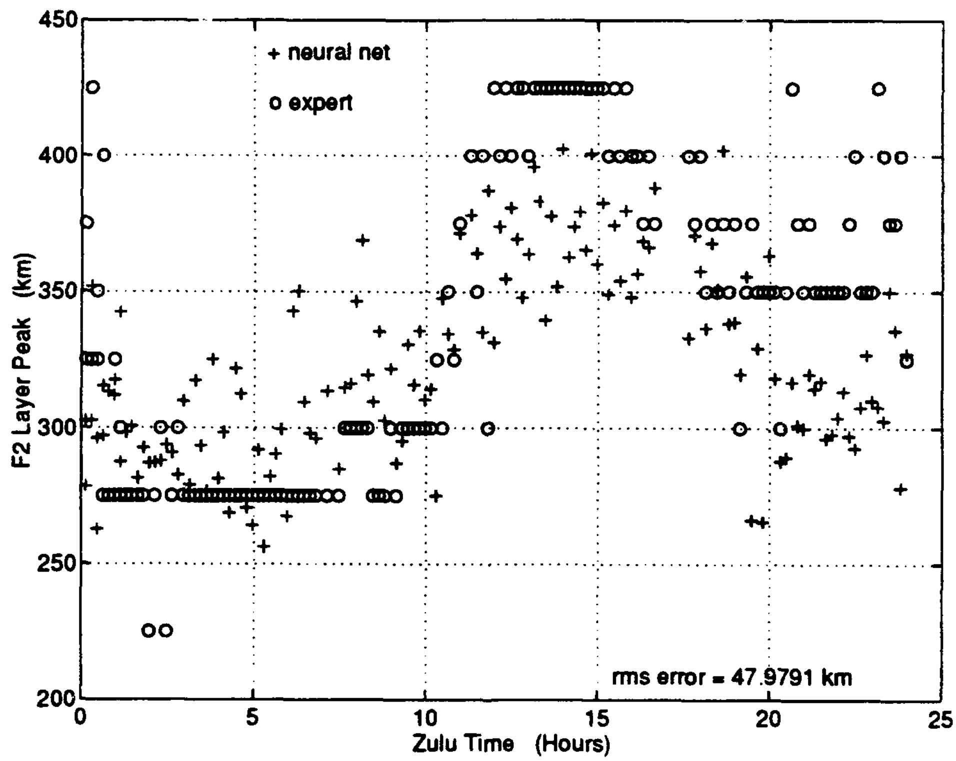

Figure 4.16 Test Set (F_2 Layer Peak) - Model Network. 10,000 training passes.

V. CONCLUSIONS

The experimental data neural network showed neural networks are excellent at modeling ionospheric data for a given day. The continuous nature of neural networks and their ability to interpolate provide for more accurate modeling than is possible when using discrete data. The neural network was good at mastering the diurnal variations of the ionosphere and all general trends were predicted.

It was shown that individual exceptions in the train set can influence the network's output. Therefore, to teach a network the best general trend it is essential to remove anomalous data from the train set.

The library data network showed neural networks are capable of learning many different ionospheric models. The network agreed well with the E layer and F_2 layer experimental data. One interpretation of this may be that for those two layers the models are very good and the experimental data is accurate.

The library data network's F_1 layer performance showed the correct diurnal variation pattern but the disappearance of the layer at night was not predicted. One possible source of this error might have been a lack training examples like the measured data.

The library data network's F_2 layer peak performance showed a correct general trend but the network's output data was quite scattered. There are two factors that may be contributing to the error; a large uncertainty in the expert data, and the modeling of the F_2 layer peak may not be quite adequate.

This thesis has shown neural networks have tremendous potential in the field of ionospheric modeling in general and ROTHR modeling in particular. Further research in this area should be made. The development of a network that has been trained on data taken during different seasons should be investigated. That could lead to the development of a universal ionosphere neural network that would be provide a single continuous model of the ionosphere.

LIST OF REFERENCES

Beale, R., and Jackson, T., *Neural Computing: An Introduction*, Adam Hilger, 1990.

Beer, T., *The Aerospace Environment*, Wykeham Publications (London), Ltd., 1976.

Caudill, M., and Butler, C., *Understanding Neural Networks: Computer Explorations*, Volume 1, Basic Networks, The MIT Press, 1993.

Craig, R. A., *The Edge of Space: Exploring the Upper Atmosphere*, Doubleday & Company, Inc., 1968.

Craig, R. A., *The Upper Atmosphere: Meteorology and Physics*, Academic Press, Inc., 1965.

GAO, *OVER-THE-HORIZON RADAR: Better Justification Needed for DOD Systems' Expansion*, U.S. General Accounting Office, 1991.

Ivanov-Kholodny, G. S., and Mikhailov, A. V., *The Prediction of Ionospheric Conditions*, D. Reidel Publishing Company, 1986.

The MathWorks, Inc., *MATLAB User's Guide*, 1992.

NeuralWare, Inc., *Neural Computing*, 1993.

NeuralWare, Inc., *Reference Guide*, 1993.

NeuralWare, Inc., *Using NeuralWorks*, 1993.

Ratcliffe, J. A., *An Introduction to the Ionosphere and Magnetosphere*, Cambridge University Press, 1972.

Sun Microsystems, Inc., *Sun System User's Guide*, 1991.

Thome, G., *Letter to Professor D. J. Collins*, 27 August 1991.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   Cameron Station
   Alexandria, Virginia 22304-6145

2. Library, Code 52 2
   Naval Postgraduate School
   Monterey, California 93943-5002

3. Chairman, Code AA 1
   Department of Aeronautics and Astronautics
   Naval Postgraduate School
   Monterey, California 93943-5000

4. D. J. Collins, Code AA/Co 2
   Department of Aeronautics and Astronautics
   Naval Postgraduate School
   Monterey, California 93943-5000

5. Joe Thomason 1
   NRL
   Code 5324
   Washington, D.C. 20375
   Building 56, Room 100

6. LT J. A. Pinkepank 3
   Training Officer
   USS Carl Vinson CVN-70
   FPO AP 96629-2840